

# 2D/3D级联卷积在分割CT肺动脉上的应用研究

黄绍辉<sup>①</sup> 严凯<sup>①</sup> 王博亮<sup>①</sup> 王弘轩<sup>①</sup> 王继伟<sup>\*</sup>

**摘 要** 医学影像分割是计算机辅助诊断的重要组成部分。针对CT影像的三维特性,提出了一种基于2D/3D级联卷积的Unet网络结构用来分割肺动脉。该结构相比基于传统2D卷积的方法,关联了第三维度信息,提高了分割准确度和泛化能力,相比基于传统3D卷积的方法提高了准确度和执行效率。实验对多套肺动脉增强CT数据集做了验证,分割准确率达到85.7%,高于传统2D和3D Unet网络,同时执行效率较3D Unet提高近30%,在CT影像分割上做到了效率和准确度的兼顾。

**关键词** CT影像分割 肺动脉 深度学习 2D/3D级联卷积

Doi:10.3969/j.issn.1673-7571.2019.05.002

[中图分类号] R445.3; R319 [文献标识码] A

Application Research of 2D/3D Cascade Convolution in Segmentation of CT Pulmonary Artery / HUANG Shao-hui, YAN Kai, WANG Bo-liang, et al//China Digital Medicine.-2019 14(5): 07 to 11

**Abstract** The segmentation of medical images is an important part of computer-aided diagnosis. In this paper, based on the three-dimensional characteristics of CT images, a Unet network structure based on 2D/3D cascade convolution structure is proposed to segment the pulmonary artery. Compared with the traditional 2D convolution method, the structure is associated with the third dimension information, which improves the segmentation accuracy and the ability of generalization, and improves the accuracy and execution efficiency compared with the traditional 3D convolution method. The CT datasets of multiple sets of pulmonary vascular enhancement were validated, and the segmentation accuracy rate was 85.7%, which was higher than that of traditional 2D and 3D Unet networks. At the same time, the execution efficiency was improved by about 30% compared with 3D Unet. The method make great progress of both efficiency and accuracy in CT image segmentation.

**Keywords** segmentation of CT image, pulmonary artery, deep learning, 2D/3D cascade convolution

**Fund project** The National Natural Science Foundation of China (No.61001144, 61102137, 61327001, 61671399)

**Corresponding author** Information Center of Xiamen University Affiliated Success Hospital, Xiamen 361003, Fujian Province, P.R.C.

## 1 背景

近年来,在医学领域,以深度学习为基础的计算机辅助诊断已经取得巨大突破。由于其准确度和鲁棒性较传统方法有优势,很多人工智能技术已经应用到临床。分割是医学影像处理中重要的环节,分割的准确与否直接影响最终的诊断结果。在三维影像分割中,由于数据相比传统二维图片多一个维度,导致处理难度增加。使用深度学习方法做CT影像分割,如果仅使用2D卷积,则丧失了第三维度的上下文关联,由于血管、气管、肿瘤等上下文关联强,尤其是肺部血管、动脉和静脉交织在一起,形态上差异不明显,因此二维卷积往往难以取得很好的分割效果。若使用3D卷积,参数比2D卷积增加了一个维度,参数量过大,会导致执行效率低下。针对这个问题,提出了一种2D/3D级联卷积的结构。通过这种结构,可使网络在保证分割准确度的同时更易收敛,并且大大减少参数量,提高执行效率。



**基金项目:** 国家自然科学基金(编号:61001144, 61102137, 61327001, 61671399)

\*通信作者: 厦门大学附属成功医院信息中心, 361003, 福建省厦门市思明区文园路92-96号

①厦门大学信息科学与技术学院计算机科学系, 361005, 福建省厦门市思明区曾厝垵西路1号

## 2 相关工作

2015年, Jonathan Long提出FCN<sup>[1]</sup>结构, 成功将深度学习应用在语义分割上。在医学影像方面, Olaf Ronneberger等人提出了Unet<sup>[2]</sup>网络结构, 该网络结构连接了网络低维和高维特征, 医学图像往往分割目标与背景边界模糊, 通过Unet这种高低维特征融合的方式, 可以有效增加语义分割的准确度, 得到精确的边界信息, 因此Unet这种结构被广泛应用于医学图像分割。Ahmed Abdulkadir等人将Unet拓展到三维<sup>[3]</sup>, 提出了基于3D卷积的3D Unet, 并且成功应用到三维影像分割中。Kuan-Lun Tseng等人提出了通过convLSTM关联第三维度的上下文信息的方法<sup>[4]</sup>, 并且应用到脑肿瘤分割。在传统方法中, 针对血管分割, T. Buelow等人提出了一种通过人工下种子点向前传播来分割肺部动脉的方法<sup>[5]</sup>, 但无法做到全自动分割。上述方法在各自研究领域上都取得了一定的成就, 但对于肺动脉的分割, 分割准确度或执行效率无法得到兼顾。

## 3 方法

**3.1 简介** 针对上述方法存在的问题, 提出了一种2D/3D级联卷积结构, 最终应用到Unet网络当中。相比Thomas Buelow提出的肺动脉分割方法, 无需人工干预, 做到全自动分割。

实验首先对输入数据进行一系列预处理, 输入网络, 获得预测概率分布, 通过阈值筛选获得分割二值图, 详细流程见图1。

**3.2 数据预处理** 算法采用肺部动脉期CT作为数据集, 预处理步骤如下。

由于不同CT数据层厚不同, 从0.5~2 mm不等。为了规范化样本, 提高准确度, 采用双线性插值法, 将

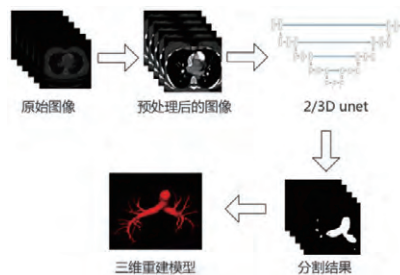


图1 肺动脉分割整体流程图

层厚统一标准化至1 mm。经标准化处理, 10套测试集CT平均分割结果准确率提高0.7%。测试集结果对比见图2。

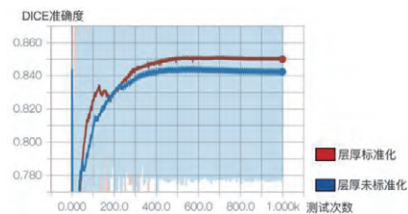


图2 层厚标准化和未标准化DICE

准确率折线图

将图像灰度值使用max-min标准化方法映射到[0,1], 归一化方法见公式(1):

$$f(x) = \frac{x - \min}{\max - \min} \quad (1)$$

其中 $x$ 为当前像素点灰度值,  $f(x)$ 为归一化后的灰度值,  $\min$ 为CT图像中灰度值最小值,  $\max$ 为灰度值最大值。

与Buelow<sup>[5]</sup>等人分割肺动脉预处理方法一样, 首先提取肺部区域。可减少需要推理的数据量, 能有效提高执行效率。同时, 由于分割内容集中在肺部, 也可避免错分肺部外组织导致的误分割。肺部分割具体算法流程见图3, 首先利用OTSU方法确定分割阈值, 其次通过区域增长方法剔除体外非肺部区域, 最后剩下的最大联通分量即为肺部区域。

以三维滑窗的方式固定采样CT图像作为网络输入, 采样大小为 $16 \times 256 \times 256$ 。采样方法为: 固定大小为 $16 \times 256 \times 256$ 的采样窗口, 训练集在x、y、z维度分别以步长为128、

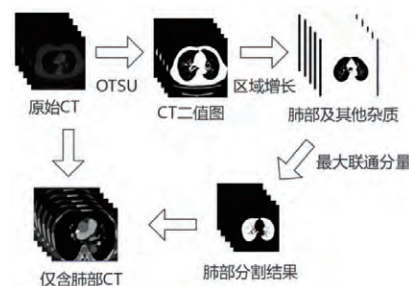


图3 肺部提取算法流程图

128、4的大小进行滑窗采样, 测试集在x、y、z维度分别以步长为256、256、16的大小进行滑窗采样, 最终获得若干数据块。训练集平均一套CT可以分割为171个数据块。测试集平均一套CT可以分割为46个数据块。

**3.3 网络结构** 算法模型以Unet模型为基础, 添加了以下改进。

**3.3.1 添加2D/3D级联卷积** 2D/3D级联结构如图4所示。

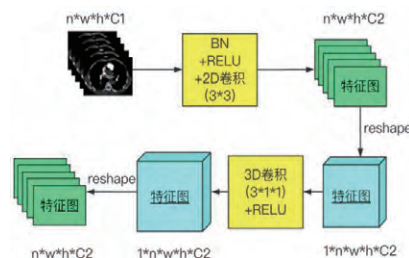


图4 2D/3D级联卷积结构

其中 $n$ 为输入切片层数, 即batch size,  $C1$ 为输入通道数,  $C2$ 为输出通道数。2D/3D级联卷积的第一层结构是卷积核大小为 $3 \times 3$ 的2D卷积, 第二层结构是卷积核大小为 $3 \times 1 \times 1$ 尺寸的3D卷积。

将多张CT切片的2D卷积结果作为3D卷积的输入, 2D卷积主要负责提取影像的二维特征, 3D卷积负责关联CT切片之间上下文特征, 这样的级联结构比2D卷积多提取到第三维特征, 相对于3D卷积, 则有效减少了需要学习的参数量, 且提高了执行效率。同时算法泛化能力也得到了提升。

**3.3.2 添加批标准化层 (Batch**



Normalization) 随着网络结构深度的增加, 网络变得难以收敛。2015年, Sergey Ioffe等人提出了批标准化层<sup>[6]</sup>, 在一定程度上解决了梯度弥散问题, 提高了收敛速度, 增加了网络的稳定性。和目前很多网络结构一样, 本网络模型添加了BN层, BN具体实现过程如下:

输入:  $B = \{x_{1...m}\}$ , 其中  $x_i$  代表batch中的每个切片的特征图。需要训练的参数  $\gamma, \beta$ 。

输出:  $\{y_i = BN_{\gamma, \beta}(x_i)\}$ 。  
 $\mu_B \leftarrow \frac{1}{m} \sum_{i=1}^m x_i$ , 计算  $x_i$  均值。

$\sigma_B^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2$ , 计算  $x_i$  方差。

$\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$ , 通过均值方差将  $x_i$  标准化。

$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv BN_{\gamma, \beta}(x_i)$ , 通过学习的参数  $\gamma, \beta$ 。恢复  $x_i$  尺寸, 避免限制在 $[-1, 1]$ 区间, 结果记为  $y_i$ 。

**3.3.3 使用Focal损失函数** 在肺部CT影像中, 一般以体素数量来计算实际体积, 动脉血管所占整体CT的比例往往小于1%, 正负样本比例严重失衡。若仅使用交叉熵作为损失函数, 会导致算法不收敛, 使模型陷入局部最优解。针对这种正负样本不均问题, Tsung-Yi Lin等人提出了Focal损失函数<sup>[7]</sup>, Focal损失函数不仅解决了正负样本不平衡的问题, 同时可让模型更加关注难样本的区分, 适用于肺血管分割的场景。Focal损失函数见公式(2):

$$FL(p_t) = -\alpha_t(1-p_t)^\gamma \log(p_t) \quad (2)$$

$$p_t = \begin{cases} p, & \text{if } \gamma = 1 \\ 1-p, & \text{otherwise} \end{cases}$$

$\alpha_t$  和  $\gamma$  为可调节参数, 采用

$\alpha_t = 0.25, \gamma = 2$ 。

**3.3.4 网络结构** 如图5所示, 2D/3D Unet在Unet的基础上, 在提取特征即编码阶段, 对于每一个尺寸的分辨率的卷积组, 第一层卷积换成2D/3D卷积, 并且在第二层卷积前添加BN层, 在解码阶段对于每一个尺寸的分辨率的卷积组, 将第一层卷积替换成2D/3D卷积, 第二层2D卷积前增加BN层。

**3.4 训练和测试** 训练部分主要分为以下几个步骤: (1) 设置Adadelta作为模型训练优化器, 初始学习率设置为0.1, 每训练500次衰减一次, 衰减率为0.98; (2) 按照3.2步骤预处理, 得到若干 $16 \times 256 \times 256$ 的数据块作为训练输入数据; (3) 为了提高模型的泛化能力, 对数据进行一定比例的数据增广, 具体方法是将CT数据以20%的概率随机旋转 $0 \sim 90^\circ$ ; (4) 将(3)得到的数据和标签送入网络, 进行训练。

测试部分主要分为以下几个步骤: (1) 设置阈值T, 设置T为0.6; (2) 按照3.2步骤预处理, 得到若干 $16 \times 256 \times 256$ 数据块, 并且记录每个数据块在CT数据中的位置; (3) 将(2)得到的数据送入网络, 进行推理, 得到概率预测图; (4) 将获得的多个数据块推理预测结果按照记录

位置, 拼接成整个CT数据的预测概率图; (5) 将概率图值P与阈值T相比, 若P大于T则为目标, P小于T则为背景。最终得到分割二值图。

## 4 实验结果

数据集为35套肺部CT数据, 该数据集来自厦门大学附属成功医院, 层厚 $0.5 \sim 2$  mm, 部分CT有肺部疾病, 标签数据由人工标注获得。数据集分为训练集和测试集, 其中训练集为25套, 测试集为10套。

实验环境操作系统为Ubuntu16.04 LST 64 位, 内存大小为64G, 处理器为Intel(R) Xeon(R) CPU E5-2690 v4 @ 2.60GHz, 显卡为NVIDIA TESLA V100 16GB, 采用tensorflow深度学习框架。

实验中通过提出的2D/3D级联卷积, 有效提高了肺动脉分割的准确度和泛化能力, 取得了较好的分割结果, 具体如图6所示。

其中第一列为原始CT切片, 第二列为原始CT切片加实际分割结果, 第三列为实际分割结果二值图, 第四列为期望标签二值图。从图6可以看到, 分割结果与期望标签差别不大, 有较高的分割准确度。

实验通过利用视觉化工具函数库

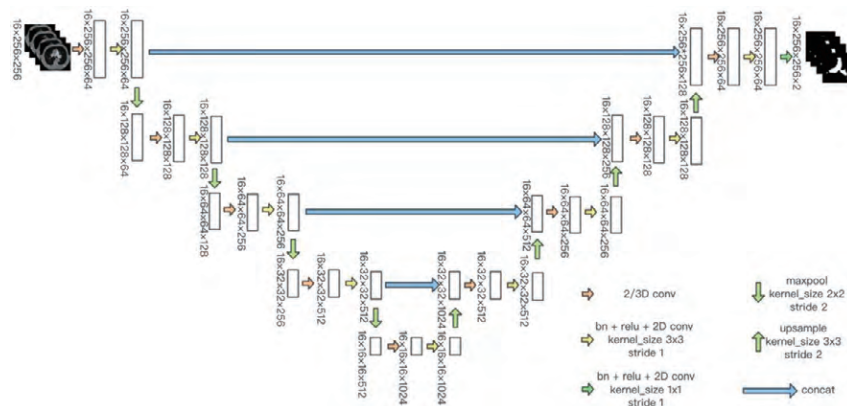


图5 2D/3D Unet结构

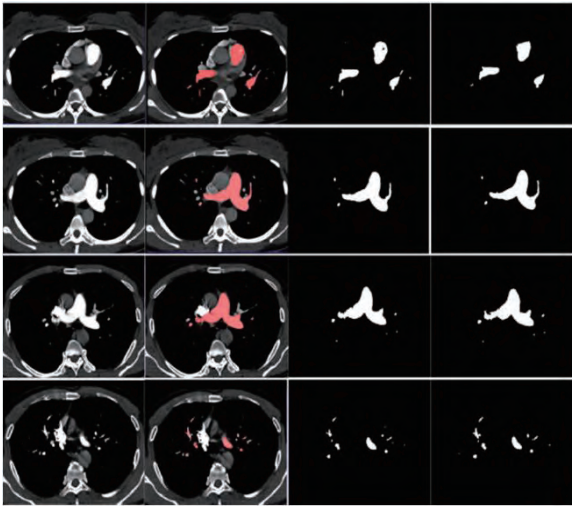


图6 测试集肺动脉分割结果

(VTK) 和医学图像分割与配准算法研发平台 (ITK) 将分割结果进行三维建模。三维建模之后的立体结构如图7所示。

其中第一列表示实际分割结果的三维模型，第二列为标签三维建模结果。

为了证明提出的2D/3D级联卷积的有效性，采用DICE体积重叠误差 (VOE)、相对体积误差 (RVD) 作为评价指标<sup>[8-9]</sup>。指标的具体计算方式如下。

体积重叠误差 (VOE)：实际分割结果体积与真实标签分割体积的重叠误差，该值越小表明分割的结果越准确。A代表预测结果，B代表实际标签，计算方法见公式 (3)：

$$VOE(A, B) = 1 - \frac{|A \cap B|}{|A \cup B|} \quad (3)$$

重合率(DICE)：该值的区间为 [0,1]，值越大，表示分割结果与实际标签越接近。计算方法见公式 (4)：

$$DICE(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (4)$$

相对体积误差 (RVD)：这个指标用于判断实际分割结果是否过分割或欠分割，计算方法见公式 (5)：

$$RVD(A, B) = \frac{|B| - |A|}{|A|} \quad (5)$$

实验通过分别对三种网络模型 (2D Unet、3D Unet、2D/3D Unet) 进行测试以此来证明2D/3D级联卷积

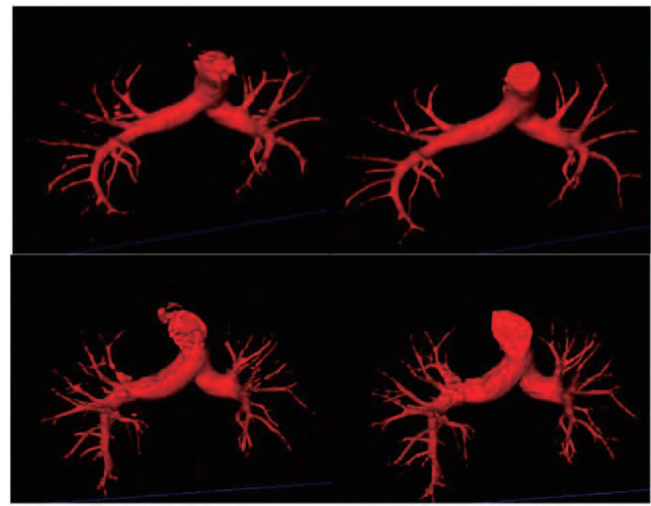


图7 测试集肺动脉三维重建结果

结构的有效性和高效性。在同样的运行环境和训练方法下实验结果见表1和表2，其中执行时间为每个数据块 (尺寸为  $16 \times 256 \times 256$ ) 的训练时间或推理时间。

从训练结果来看，2D/3D Unet 和 2D Unet在拟合程度上差距不大，都有较好的拟合能力，且执行效率高。相比之下3D Unet拟合能力不如前两者，且执行效率较低。

从以上测试结果可以看出，提出的2D/3D结构在肺动脉分割的应用上，相比2D Unet在执行时间相差不大的情况下，准确度上有大幅度的提升，泛化能力更强。比3D Unet虽然准确度上提升不明显，但是执行效率上有明显的优势。

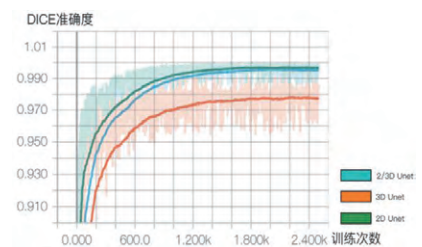


图8 训练集DICE准确度折线图

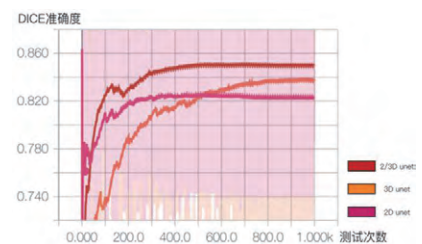


图9 测试集DICE准确度折线图

从图8、图9中可以看出，2D/3D Unet不仅在泛化能力上高于其他两种

表1 训练集肺动脉分割结果对比

MODEL	DICE	VOE	RVD	COST_TIME/s
2D Unet	99.69%	0.65%	0.16%	0.2953
3D Unet	97.71%	4.46%	1.48%	0.8021
2D/3D Unet	99.57 %	0.85%	0.24%	0.4161

表2 测试集肺动脉分割结果对比

MODEL	DICE	VOE	RVD	COST_TIME/s
2D Unet	82.74%	28.69%	13.93%	0.1109
3D Unet	85.16%	25.56%	14.3%	0.2251
2D/3D Unet	85.77%	24.61%	9.6%	0.1498



模型,同时在模型收敛速度上也有一定的优势。

## 5 总结

本文提出了一种基于2D卷积和3D卷积级联的结构,将不同切片的二维特征融合,达到提取三维特征的目的。将此结构应用到Unet网络结构当中,提出了一种改良模型2D/3D Unet。在分割肺动脉的实验中,与传统2D Unet和3D Unet作对比,2D/3D Unet在准确度得到提升的同时,执行效率也得到了保证,充分证明了2D/3D结构的有效性和高效性。除了肺动脉,对于CT数据肝脏的分割也能获得良好的分割结果。由于篇幅原因未做具体阐述。

另外,该结构有很好的移植性,可以简单应用到其他网络结构。下一步的工作内容在于验证2D/3D级联结构对于除了Unet以外的分割模型是否同样有效。

## 参考文献

- [1] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation[C]. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 3431-3440.
- [2] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[J]. MICCAI, 2015, 9351: 234-241.
- [3] Çiçek Ö, Abdulkadir A, Lienkamp SS. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation[C]. Medical Image Computing and Computer-Assisted Intervention, MICCAI 2016.
- [4] Kuan-Lun Tseng, Yen-Liang Lin, Winston Hsu. Joint Sequence Learning and Cross-Modality Convolution for 3D Biomedical Segmentation[C]. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 6393-6400.
- [5] Buelow T, Wiemker R, Blaffert T, et al.

- Automatic extraction of the pulmonary artery tree from multi-slice CT data[J]. Proc. SPIE: Med. Imag., 2005(5746): 730-740.
- [6] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]. In ICML, 2015.
- [7] Lin TY, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. arXiv preprint arXiv, 2017(1708): 02002.
- [8] 郭树旭, 马树志, 李晶, 等. 基于全卷积神经网络的肝脏CT影像分割研究[J]. 计算机工程与应用, 2017, 53(18): 126-131.
- [9] Zhang Y, He Z, Zhong C, et al. Fully convolutional neural network with post-processing methods for automatic liver segmentation from CT[C]. Proceedings of Chinese Automation Congress. Washington DC, USA: IEEE Press, 2017: 3864-3869.

【收稿日期: 2019-03-08】

(责任编辑: 肖婧婧)

(上接第06页) 隐私安全评估的迫切性不言而喻, 评价既是准入, 也是监管, 识别平台在数据开放共享过程中的风险点, 提前做好预防措施, 降低泄漏风险, 切实推进医疗数据的开放共享, 释放数据红利。

## 参考文献

- [1] 陈敏, 刘宁. 健康医疗大数据发展现状研究[J]. 医学信息学杂志, 2017, 38(7): 2-6.
- [2] 国务院发布《关于促进和规范健康医疗大数据应用发展的指导意见》[EB/OL]. [2019-03-10]. [http://www.gov.cn/zhengce/content/2016-06/24/content\\_5085091.htm](http://www.gov.cn/zhengce/content/2016-06/24/content_5085091.htm).
- [3] 黄匡时. 我国健康医疗大数据的现状、问题及对策[J]. 中国数字医学, 2017, 12(5): 24-26.

- [4] 国信办印发《信息安全风险评估指南》[EB/OL]. [2019-03-10]. <http://www.doc88.com/p-1405049786157.html>.
- [5] United Kingdom: Data protection act 1998. [EB/OL]. [2019-03-10]. [http://legislation.gov.uk/ukpga/1998/29\(1998\)](http://legislation.gov.uk/ukpga/1998/29(1998)).
- [6] Personal Information Protection Law in Japan [EB/OL]. [2019-03-10]. <https://www.jonesday.com/Personal-Information-Protection-Law-in-Japan-11-09-2005/>.
- [7] European Commission: Proposal for a regulation of the European parliament and of the council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (general data protection regulation) [EB/OL]. [2019-03-

- 10]. [http://europa.eu/rapid/press-release\\_IP-16-3674\\_en.htm](http://europa.eu/rapid/press-release_IP-16-3674_en.htm). (2000).
- [8] 马诗诗, 于广军, 崔文彬. 患者医疗信息隐私保护现状与需求调查[J]. 中华医院管理杂志, 2018, 34(1): 55-58.
- [9] Grime MM, Wright G. Delphi Method[M]. Wiley StatsRef: Statistics Reference Online, 2016.
- [10] 王晓丽, 孙雪松, 张寅, 等. 区域卫生信息管理平台指标体系构建研究[J]. 中国数字医学, 2017, 12(1): 19-21, 28.

【收稿日期: 2019-03-25】

【修回日期: 2019-05-05】

(责任编辑: 肖婧婧)